

# **PENERAPAN DATA MINING UNTUK MENGETAHUI POLA LAMA STUDI MAHASISWA MENGGUNAKAN ALGORITMA C4.5**

(Studi kasus : Program Studi Teknik Informatika UNPAS Bandung)

## **TUGAS AKHIR**

Disusun sebagai salah satu syarat untuk kelulusan Program Strata 1,  
di Program Studi Teknik Informatika, Universitas Pasundan Bandung

oleh :

Shelly Yolanda  
NRP : 12.304.0229



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS PASUNDAN BANDUNG  
JANUARI 2017**

## DAFTAR ISI

|   |      |
|---|------|
| ABSTRAK .....   | i    |
| ABSTRACT.....   | ii   |
| KATA PENGANTAR .....  | iii  |
| DAFTAR ISI .....  | iv   |
| DAFTAR TABEL .....  | vi   |
| DAFTAR GAMBAR .....   | vii  |
| BAB 1 PENDAHULUAN.....  | 1-1  |
| 1.1 Latar Belakang Masalah .....                                  | 1-1  |
| 1.2 Identifikasi Masalah .....                                    | 1-2  |
| 1.3 Tujuan Tugas Akhir .....                                      | 1-2  |
| 1.4 Lingkup Tugas Akhir .....                                     | 1-3  |
| 1.5 Metodologi Tugas Akhir.....                                   | 1-3  |
| 1.6 Langkah-langkah Pengerjaan.....                               | 1-4  |
| 1.7 Sistematika Penulisan Laporan.....                            | 1-5  |
| BAB 2 LANDASAN TEORI.....   | 2-1  |
| 2.1. Knowledge Discovery In Database (KDD) .....                  | 2-1  |
| 2.2. Data Mining.....   | 2-2  |
| 2.2.1. Pengertian Data Mining.....                                | 2-2  |
| 2.2.2. Permasalahan yang Memotivasi Pengembangan Data Mining..... | 2-3  |
| 2.2.3. Educational Data Mining.....                               | 2-3  |
| 2.2.4. Pengenalan Pola.....                                       | 2-4  |
| 2.2.5. Pengelompokan Data Mining.....                             | 2-4  |
| 2.2.6. Konsep Klasifikasi.....                                    | 2-5  |
| 2.2.7. Pengukuran Kinerja Klasifikasi.....                        | 2-6  |
| 2.2.8. Pohon Keputusan (Decision Tree) .....                      | 2-7  |
| 2.2.9. Algoritma C4.5.....  | 2-8  |
| 2.3. K-Fold Cross Validation.....                                 | 2-9  |
| 2.4. RapidMiner.....  | 2-10 |
| 2.5. Penelitian Terdahulu.....                                    | 2-10 |
| BAB 3 SKEMA PENELITIAN.....                                       | 3-1  |
| 3.1. Rancangan Penelitian.....                                    | 3-1  |

|   |            |
|---|------------|
| 3.2. Peta Analisis.....   | 3-2        |
| 3.3. Analisis.....  | 3-3        |
| 3.3.1. Analisis Data Mining untuk Mengetahui Pola Lama Studi Mahasiswa..... | 3-3        |
| 3.3.2. Analisis Kerangka Pemikiran.....                                     | 3-4        |
| 3.3.3. Analisis Kebutuhan Data.....   | 3-5        |
| 3.3.4. Analisis Kakas Data Mining.....                                      | 3-6        |
| <b>BAB 4 IMPLEMENTASI.....</b>  | <b>4-1</b> |
| 4.1. Persiapan.....   | 4-1        |
| 4.1.3. Instalasi RapidMiner.....  | 4-1        |
| 4.1.4. Tahapan Knowledge Discovery In Databases (KDD) .....                 | 4-2        |
| 4.2. Tahapan Data Mining Menggunakan Perhitungan Manual.....                | 4-8        |
| 4.3. Tahapan Data Mining Menggunakan Aplikasi RapidMiner.....               | 4-28       |
| 4.3.1. Fase Pembentukan Model.....  | 4-28       |
| 4.3.2. Fase Pengujian Model.....  | 4-37       |
| 4.3.3. Perbandingan Model yang Dihasilkan.....                              | 4-41       |
| 4.3.4. Penggunaan Model.....  | 4-41       |
| <b>BAB 5 KESIMPULAN DAN SARAN.....</b>                                      | <b>5-1</b> |
| 5.1. Kesimpulan.....  | 5-1        |
| 5.2. Saran.....   | 5-1        |
| <b>DAFTAR PUSTAKA.....</b>  | <b>ix</b>  |

## DAFTAR TABEL

|   |      |
|---|------|
| Tabel 2. 1 Matriks konfusi untuk klasifikasi dua kelas [PRA12]..... | 2-6  |
| Tabel 2. 2 Penelitian Terdahulu.....                                | 2-10 |
| Tabel 3. 1 Langkah Penyelesaian Penelitian.....                     | 3-1  |
| Tabel 3. 2 Detail Analisis.....                                     | 3-3  |
| Tabel 3. 3 Hasil Wawancara.....                                     | 3-5  |
| Tabel 4. 1 Spesifikasi Kebutuhan Hardware.....                      | 4-1  |
| Tabel 4. 2 Spesifikasi Kebutuhan Software.....                      | 4-1  |
| Tabel 4. 3 Data Mahasiswa.....                                      | 4-3  |
| Tabel 4. 4 Data Asisten.....  | 4-4  |
| Tabel 4. 5 Hasil Integrasi.....                                     | 4-4  |
| Tabel 4. 6 Proses Pembersihan Data.....                             | 4-5  |
| Tabel 4. 7 Hasil Pembersihan Data.....                              | 4-5  |
| Tabel 4. 8 Transformasi Data.....                                   | 4-6  |
| Tabel 4. 9 Hasil Transformasi Data.....                             | 4-7  |
| Tabel 4. 10 Dataset.....  | 4-8  |
| Tabel 4. 11 Perhitungan Iterasi 0.....                              | 4-13 |
| Tabel 4. 12 Data Sorting Berdasarkan IPS3.....                      | 4-15 |
| Tabel 4. 13 Perhitungan iterasi1.....                               | 4-19 |
| Tabel 4. 14 Data Sorting Berdasarkan IPS2.....                      | 4-21 |
| Tabel 4. 15 Perhitungan iterasi2.....                               | 4-25 |
| Tabel 4. 16 Hasil Pengujian Model.....                              | 4-28 |
| Tabel 4. 17 Hasil Perbandingan Model.....                           | 4-41 |
| Tabel 4. 18 Contoh Data yg belum diketahui kelasnya.....            | 4-41 |
| Tabel 4. 19 Data baru yang sudah diketahui kelasnya.....            | 4-42 |

## DAFTAR GAMBAR

|  |      |
|--|------|
| Gambar 2. 1 Proses Dari Penemuan Pengetahuan [HAN12] .....       | 2-1  |
| Gambar 2. 2 Proses Pekerjaan Klasifikasi [PRA12] .....           | 2-6  |
| Gambar 2. 3 Gambar Metode K-Fold Cross Validation [OLS08] .....  | 2-9  |
| Gambar 3. 1 Contoh Pohon Keputusan [PRA07] .....                 | 2-7  |
| Gambar 3. 2 Peta Analisis.....                                   | 3-2  |
| Gambar 3. 3 Analisis Pohon Masalah.....                          | 3-3  |
| Gambar 3. 4 Konsep Pembentukan Model.....                        | 3-4  |
| Gambar 3. 5 Konsep Pengujian Model.....                          | 3-4  |
| Gambar 4. 1 Pohon Perhitungan Iterasi 0.....                     | 4-15 |
| Gambar 4. 2 Pohon Keputusan Perhitungan Iterasi 1.....           | 4-20 |
| Gambar 4. 3 Pohon Keputusan Hasil Iterasi 2.....                 | 4-26 |
| Gambar 4. 4 Pohon Keputusan dengan Tools.....                    | 4-27 |
| Gambar 4. 5 Tampilan Awal RapidMiner.....                        | 4-28 |
| Gambar 4. 6 Tampilan Import Data.....                            | 4-29 |
| Gambar 4. 7 Tampilan Import Data (2) .....                       | 4-30 |
| Gambar 4. 8 Tampilan Import Data (3) .....                       | 4-30 |
| Gambar 4. 9 Tampilan Proses Mining.....                          | 4-31 |
| Gambar 4. 10 Tampilan Proses Mining (2) .....                    | 4-31 |
| Gambar 4. 11 Tampilan Hasil Berupa Pohon Keputusan.....          | 4-32 |
| Gambar 4. 12 Akar (root) .....                                   | 4-32 |
| Gambar 4. 13 Node Keputusan IPS4 = Cum Laude.....                | 4-33 |
| Gambar 4. 14 Node Keputusan IPS4=Kurang Memuaskan.....           | 4-33 |
| Gambar 4. 15 Node Keputusan IPS4=Memuaskan.....                  | 4-34 |
| Gambar 4. 16 Node Keputusan IPS4=Sangat Memuaskan.....           | 4-34 |
| Gambar 4. 17 Hasil dalam bentuk Deskripsi.....                   | 4-37 |
| Gambar 4. 18 Tampilan Proses Pengukuran Kinerja Klasifikasi..... | 4-38 |
| Gambar 4. 19 Hasil Pengukuran Kinerja Klasifikasi.....           | 4-38 |
| Gambar 4. 20 Tampilan Proses Evaluasi.....                       | 4-39 |
| Gambar 4. 21 Tampilan Proses Evaluasi (2) .....                  | 4-40 |
| Gambar 4. 22 Hasil Evaluasi Model.....                           | 4-40 |
| Gambar 4. 23 Tampilan Pohon Keputusan.....                       | 4-42 |